

文章编号: 1000-5862(2018)05-0441-06

基于可达阵的补偿模型 Q 矩阵标定方法

汪文义¹ 汪 腾¹ 宋丽红² 高 朋¹

(1. 江西师范大学计算机信息工程学院 江西 南昌 330022; 2. 江西师范大学初等教育学院 江西 南昌 330022)

摘要: Q 矩阵标定是认知诊断评估中研究的热点问题, Q 矩阵的好坏决定了认知诊断评估的准确性. 根据确定性输入噪声“与”门模型(DINA)中可达阵 R 与简化 Q 矩阵存在布尔“与”的关系, 提出基于确定性输入噪声“或”门模型(DINO)的可达阵 R 与简化 Q 矩阵在列向量上存在布尔“或”的关系, 并由此推导出基于可达阵的补偿模型 Q 矩阵标定方法. 实验结果表明: 当可达阵失误与猜测小于 0.20 且待标定项目参数小于 0.25 时, 该方法所得 Q 矩阵元素返真率达到 90% 以上, 且在可达阵失误与猜测参数均小于 0.25 时真实 Q 矩阵与估计 Q 矩阵之间的差异较小.

关键词: 认知诊断评估; 可达阵; 补偿模型; Q 矩阵标定

中图分类号: B 841 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2018.05.01

0 引言

传统的教学测验只能通过分数的高低将学生进行分类, 对他们知识点的掌握情况无法给出具体的信息. 我国《国家中长期教育改革和发展规划纲要(2010—2020年)》强调因材施教, 关注学生不同特点和个性差异, 发展每一个学生的优势潜能. 2016年教育部印发《教育信息化“十三五”规划》, 促进教育信息化全面深入应用, 使教学更加个性化. 要实现个性化教育和个性化学习, 均离不开诊断测试.

早在2001年时任美国总统的布什提出了“不让一个孩子掉队(No Child Left Behind)法案”, 2009年美国总统奥巴马提出的“力争上游(Race to the Top)”教改计划, 均要求测验除了提供测验结果外, 还要求给教师提供诊断性信息, 这时认知诊断评估的作用就充分显现出来. 通过认知诊断评估, 以及对测验作答反应分析, 教师和学生能够及时了解测验所考察的属性或知识技能掌握情况, 为以后的教学质量和学生成绩的提高提供帮助.

认知诊断评估主要用于测量被试的知识结构和加工技能(简称属性). 认知诊断评估的终极目标, 是服务于学习和学习进程的评估. 构建 Q 矩阵, 是

实现认知诊断评估十分关键的步骤之一. Q 矩阵描述了测验项目与属性之间的关联关系, 其元素 $q_{jk} = 1$ 表示项目 j 考查了属性 k , $q_{jk} = 0$ 表示项目 j 没有考查属性 k . Q 矩阵及其正确性直接决定测验的信度和效度. 国内外对于属性间无补偿关系模型的 Q 矩阵标定方法的研究已经相当成熟, 特别是确定性输入噪声“与”门(deterministic inputs, noisy "and" gate model, DINA)模型. Q 矩阵标定方法主要有: δ 方法^[1]、 γ 方法^[2]、贝叶斯估计方法^[3]、数据驱动方法^[4-5]、最小残差平方和方法^[6]、在线标定方法及其推广方法^[7-10].

虽然许多学者也研究属性间存在补偿关系的认知诊断模型以及应用的场景, 例如, E. Maris^[11] 在研究 Tatsuoka 分数加法问题时提出对于有多种解题策略的项目, 只需掌握任一种解答策略就可以有较高概率正确作答该项目, 并总结出了补偿型分类潜在类别模型(Compensatory multiple classification latent class model, MCLCM-C). J. L. Templin 等^[12] 在研究《精神疾病诊断与统计手册》第4版中病理性赌博的诊断指标(视为属性)时, 提出的确定性输入噪声“或”门(deterministic inputs, noisy "or" gate model, DINO)模型, 只要被试者具有项目所测的一方面障碍, 就会对赌博研究量表中的项目给出正向

收稿日期: 2018-05-13

基金项目: 国家自然科学基金(31500909, 31360237, 31160203, 30860084), 全国教育科学规划教育部重点课题(DHA150285)和江西省自然科学基金(20161BAB212044)资助项目.

作者简介: 汪文义(1983-), 男, 湖南衡山人, 副教授, 博士, 主要从事教育测量与信息处理的研究. E-mail: wenyiwang@jxnu.edu.cn

作答.但是这些应用中的测验 Q 矩阵属性向量都是事后找专家标定.特别是对于存在多解题策略的题目,专家需要标定多个测量模式.这样要花费巨大的开销而且专家标定存在人为主观因素的干扰,无法反应出被试者的真实心理结构.因此,詹沛达等^[13]提议在分析已有心理学量表并构建 Q 矩阵时可尝试采用基于因素分析技术的 Q 矩阵或采用事前分析的方式去重新编制类似于属性层级模型中 Q 矩阵的构建与项目编制的心理学量表.

本文在丁树良等提出的 Q 矩阵理论之扩张算法^[14]、累赘表达式的分解^[15]和 DINA 模型下 Q 矩阵标定方法^[16]的基础上,当专家事先标定少量测验试题的属性向量,且可达阵 R 与简化 Q 矩阵存在布尔“或”关系时,提出基于可达阵的补偿模型 Q 矩阵标定方法,将未标定的测验试题属性向量标定,并验证该方法的正确性.模拟分析中将采用 0-1 计分方式 DINO 模型进行研究.

1 Q 矩阵标定方法

1.1 确定性输入噪声“或”门模型

与 DINA 模型这类属性离散的非补偿模型不同的是,DINO 模型是一种经典的属性离散的补偿模型.DINO 模型是 2006 年由 Templin 等在使用认知诊断模型进行传统心理学量表的数据分析所提出的一种新模型.其目的在于探究一些隐藏的、有意义的信息^[13].DNIO 模型的理想反应模式定义为 $\omega_{ij} = 1 - \prod_{k=1}^K (1 - \alpha_{ik})^{q_{jk}}$ ^[12],其中 ω_{ij} 表示知识状态为 α_i 的被试在项目 j 上的潜在反应.

该模型在涉及考察 2 个或 2 个以上的属性时,属性与属性之间能够相互补偿,即被试者在掌握一个或多个属性可以弥补其对另一个或多个属性上的缺失.表现在实际作答中,若某项目只考察 2 个属性,当被试者在作答该项目时完全掌握其中的某一个属性,而对另一个属性没有完全掌握或者说掌握的很糟糕,则他也有较大概率正确作答该项目.

列举一个冬季奥林匹克运动会项目中的例子——现代冬季两项.参加这个项目的运动员需要完成 20 km 越野滑雪和 4 次射击.如果射击脱靶,那么每发子弹加罚 1 min.最终运动员的成绩是以越野滑雪所用时长加上加罚时间作为最后的总时间,通过用时的多少来决定名次.从这个例子中可以得出,项目之间是离散的,若射击百发百中,则就能补偿你在越野滑雪上的不足.同样地,若越野滑雪所用时长

较短,则即使射击环节有几发没命中,运动员也能获得一个好的名次.

1.2 基于可达阵的补偿模型 Q 矩阵标定方法

为了定理证明的表述简便,先对某些记号加以说明. K 表示属性数, $R_{K \times K}$ 为可达阵,其中 $[r_1 \ r_2 \ r_3 \ \cdots \ r_K]$ 为可达阵 $R_{K \times K}$ 的列剖分矩阵. Q_S 表示知识状态全集,其中 $\alpha_i \in Q_S$. Q_r 表示简化 Q 矩阵, q_j 为项目 j 所考察的属性向量,且 $q_j \in Q_r$. 一个测验可用属性与项目的关联矩阵表示,这个关联矩阵用 Q_i 表示.记 $Q_i = [R_{K \times K} \ q_j] = [r_1 \ r_2 \ r_3 \ \cdots \ r_K \ q_j]$, $q_j = \bigvee_{k \in S_j} r_k$,其中 S_j 表示可达阵中可以通过布尔“或”表示属性向量 q_j 的列标集合.此时 q_j 表示项目 j 的属性向量由可达阵中某列或者某几列通过布尔“或”运算得来.记 A_j 表示属性向量为 q_j 元素为 1 的行标集合,即 $A_j = \{k: q_{kj} = 1, k = 1, 2, \cdots, K\}$. 假设 K 个属性相互独立,即对于 Q_S 中含有 2^K 个不同知识状态.

定理 1 假设 K 个属性相互独立, $q_j = \bigvee_{l \in S_j} r_l$, 当且仅当 $\forall \alpha_i \in Q_S, w_{ij} = \bigvee_{l \in S_j} w_{il}$.

证 (i) 先证必要性. 对于 $\alpha_i \in Q_S$, 若 $w_{ij} = 1$, 则 $\exists k \in A_j$, 有 $1 - \alpha_{ki} = 0$ 且 $q_{kj} = 1$. 因为 $q_{kj} = \bigvee_{l \in S_j} w_{il} = 1$, 则 $\exists l \in S_j$, 有 $r_{kl} = 1$, 由 $1 - \alpha_{ki} = 0$ 可知 $w_{il} = 1$, 从而 $\bigvee_{l \in S_j} w_{il} = 1$. 因此 $q_{kj} = \bigvee_{l \in S_j} w_{il} = 1$. 对于 $\alpha_i \in Q_S$, 若 $w_{ij} = 0$, 则 $\forall k \in A_j$, 有 $1 - \alpha_{ki} = 1$ 且 $q_{kj} = 1$, 因为 $q_{kj} = \bigvee_{l \in S_j} w_{il} = 1$, 则 $\forall l \in S_j$, 无论 $r_{kl} = 0$ 或 $r_{kl} = 1$ 均有 $(1 - \alpha_{ki})^{r_{kl}} = 1$, 于是 $w_{il} = 0$, 从而 $\bigvee_{l \in S_j} w_{il} = 0$. 因此 $w_{ij} = \bigvee_{l \in S_j} w_{il} = 0$.

(ii) 再证充分性(反证法). 若 $\exists k \notin A_j$, 有 $\bigvee_{l \in S_j} r_{kl} = 1$ 且 $q_{kj} = 0$, 则 $\exists l \in S_j$, 有 $r_{kl} = 1$. 由 α_i 的任意性且 K 个属性相互独立得, 取 α_i 满足 $\alpha_{kl} = 1$ 且 $\alpha_{k'i} = 0, k' \neq k, k' = 1, 2, \cdots, K$ 或满足 $\alpha_{k'i} = 1, \forall k'' \in A_l - A_j, \alpha_{k'i} = 0, \forall k' \in A_j$ (这里 $k \in A_l - A_j$), 则 $w_{ij} = 0$. 而 $w_{il} = 1$, 进而 $\bigvee_{l \in S_j} w_{il} = 1$, 则与已知矛盾. 若 $\exists k \in A_j$, 有 $\bigvee_{l \in S_j} r_{kl} = 0$ 且 $q_{kj} = 1$, 由 α_i 的任意性且 K 个属性相互独立得, 同样取 α_i 满足 $\alpha_{ki} = 1$ 且 $\alpha_{k'i} = 0, k' \neq k, k' = 1, 2, \cdots, K$ 或满足 $\alpha_{k'i} = 1, \forall k'' \in A_l - A_j, \alpha_{k'i} = 0, \forall k' \in A_j$ (这里 $k \in A_l - A_j$), 其中 $q_{j'} = \bigvee_{l \in S_j} r_{l'}$, 则 $w_{ij} = 1$. 而 $w_{il} = 0$, 进而 $\bigvee_{l \in S_j} w_{il} = 0$, 则与已知矛盾. 因此, $\forall k$, 有 $q_{kj} = \bigvee_{l \in S_j} r_{kl}$.

$$\text{设 } Q_S = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} R =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} Q_r = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} Q_s \text{ 在 } Q_r$$

上的理想反应矩阵为

$$I_{RP} = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_7] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

因此在理想反应条件下,要正确作答后4列对应的题目,只需掌握前3题所考查的其中某一个属性.可见,约简Q矩阵上的理想反应列可以由可达阵上某些理想反应列的布尔“或”表示.因为 Q_r 第4列可由 Q_r 中第1、2列布尔“或”表示,由理想反应 I_{RP} 矩阵可知,第4列可由 I_{RP} 第1、2列布尔“或”表示.反过来也成立, I_{RP} 第5列可由 I_{RP} 第1、3列布尔“或”表示,由此可知 I_{RP} 第5列对应题目的属性向量可由 I_{RP} 中第1、3列对应题目的属性向量布尔“或”表示,这从 Q_r 可以得到印证,即 Q_r 第5列可由 Q_r 第1、3列布尔“或”表示.

下面给出基于可达阵的补偿模型Q矩阵标定方法.不妨只考虑一个未标定的试题,其属性向量为 q_j ,其中 $q_j \in Q_r$.记测验Q阵 $Q_t = [R_{K \times K} \quad q_j] = [r_1 \quad r_2 \quad \cdots \quad r_K \quad q_j]$,其对应的得分矩阵为 $U = [X \quad Y_j]$.若测验中有K个项目属性向量已经由专家标定,并且这些属性向量正好组成一个完整的R,则由观察作答反应,项目j对应的可达阵列集合估计 \hat{S}_j 及属性向量估计 \hat{q}_j 分别为

$$\hat{S}_j = \arg \min_{S \in P(\{r_1, r_2, \dots, r_K\}) - \emptyset} (Y_j - \bigvee_{r_k \in S} X_k)^T (Y_j - \bigvee_{r_k \in S} X_k), \quad (1)$$

$$\hat{q}_j = \bigvee_{k \in \hat{S}_j} r_k. \quad (2)$$

2 模拟研究

2.1 研究目的

由于在真实测验环境下,往往无法确定测验项目所考查的属性向量值,因此通过模拟研究来验证基于可达阵的补偿模型Q矩阵标定方法在各种条件下的表现.

2.2 研究设计

设相互独立的属性数目 $K = 5$.测验Q矩阵 $Q_t = [R \quad Q_r]$,其中R为可达阵(单位阵) Q_r 为31列待标定的Q矩阵.被试知识状态从 Q_s 中随机抽样.前10个样本量按30递增,即30,60,90,...,300,后4个样本量分别为500,1000,3000,5000,对总共14个样本量进行研究.待标定的项目参数水平分为0,0.05,0.10,...,0.45,为了反映可达阵的项目参数对待标定项目标定准确率的影响,将可达阵R上项目参数水平同样设置为10个水平,每种实验条件下重复模拟得分阵30次.

2.3 研究步骤

(i) 根据实验条件模拟测验Q矩阵 Q_t 、项目参数和被试者的知识状态;(ii) 将第1步得到的数据代入DINO模型中模拟 $Q_t = [R \quad Q_r]$ 上得分阵;(iii) 由R上得分阵 $X = [X_1 \quad X_2 \quad \cdots \quad X_K]$ 和 Q_r 中第j列的得分向量 Y_j (1)式估计 \hat{S}_j ,再由(2)式根据 \hat{S}_j 和Q矩阵中已知部分R计算 Q_r 中第j列的 \hat{q}_j ;(iv) 按照(iii)的方法依次估计 Q_r 中所有列,估计结果记为 $\hat{Q}_r = (\hat{q}_{jk})$.

2.4 评价指标

根据Q矩阵元素返真率(CCR)和平均属性判准确率(AMR)对本实验结果进行分析.

由估计Q矩阵 $\hat{Q}_r = (\hat{q}_{jk})$ 和真实Q矩阵 $Q_r = (q_{jk})$ 可得2个矩阵中对应元素相等的比率,即Q矩阵元素返真率(CCR):

$$C_{CR} = \frac{1}{KM} \sum_{k=1}^K \sum_{j=1}^M I(\hat{q}_{kj} = q_{kj}),$$

其中项目数 $M = 31$,I为示性函数.由于存在许多小样本情形,采用与样本量大小无关的海明距离法^[17]

估计被试者知识状态 α ,结合模拟的知识状态 α 计算平均属性判准确率(AMR),计算公式如下:

$$A_{MR} = \frac{1}{KN} \sum_{k=1}^K \sum_{i=1}^N I(\hat{\alpha}_{ik} = \alpha_{ik}).$$

为了比较真实Q矩阵和新方法标定的Q矩阵下知识状态的分类准确率,记基于估计和真实Q矩阵所得到的 A_{MR} 分别为 $A_{MR}(\hat{Q})$ 和 $A_{MR}(Q)$,并计算 $A_{MR}(Q) - A_{MR}(\hat{Q})$ 的均值.

2.5 实验结果

2.5.1 Q矩阵的元素返真率 (i) 考虑样本量对Q

矩阵的 C_{CR} 的边际影响. 图 1 给出了样本量与 Q 矩阵的 C_{CR} 之间的关系图. 从总体上来说, Q 矩阵的 C_{CR} 随着样本量的增加而逐渐递增. 样本量在 300 ~ 1 000 时的增长幅度较大, 但当可达阵上猜测与失误较大时, 例如在 (0.35, 0.45) 范围内, 样本量对 C_{CR} 影响不明显. 从图 1 中可以观察到, 当待标定的项目参数小于 0.25 时, 可达阵失误与猜测参数无论如何增大, 样本量在 150 ~ 5 000 时的增幅都不明显. (ii) 考虑可达阵或待标定项目参数对 Q 矩阵的 C_{CR} 的边际影响. 图 2 分别呈现的是当样本量为 150 时可达阵和待标定项目参数与 Q 矩阵的 C_{CR} 之间关系的折线图. 从图 2 可以看出, 随着可达阵上项目参数值的

增大 Q 矩阵的 C_{CR} 明显降低. 因此, 可以得出可达阵上项目参数对 Q 矩阵的 C_{CR} 之间的影响更大这一结论. 而当待标定的项目参数小于 0.15 时, Q 矩阵的 C_{CR} 的影响基本没有太大变化. (iii) 综合考虑上述 2 个参数对 Q 矩阵的 C_{CR} 的边际影响, 即综合考虑样本量与项目参数对标定方法的影响. 表 1 列出的是当样本量为 150 时, 可达阵与待标定项目参数下标定的 Q 矩阵的 C_{CR} . 从表 1 可以看出 Q 矩阵的 C_{CR} 随着可达阵和待标定项目参数的增大而降低. 当可达阵项目参数小于 0.20 且待标定的项目参数小于 0.25 时 Q 矩阵的 C_{CR} 大于 90%.

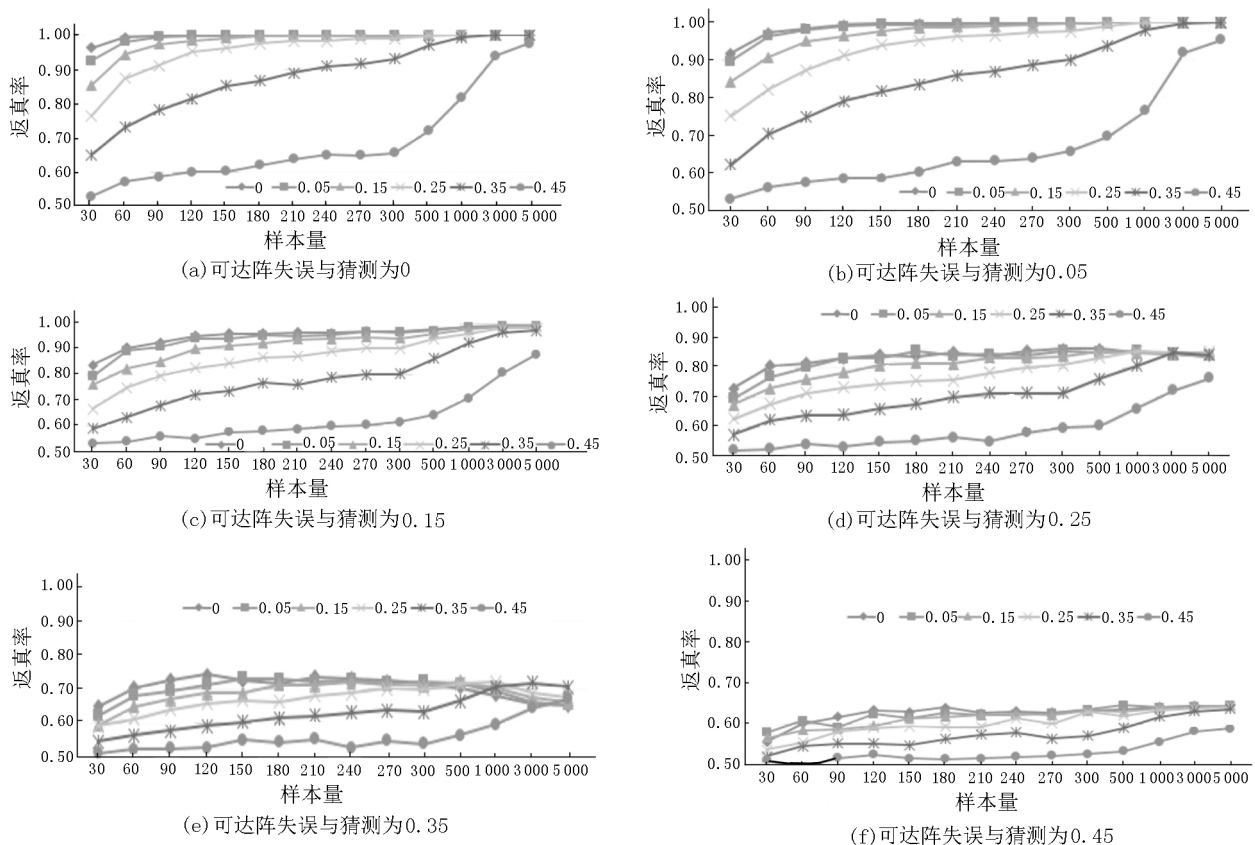


图 1 样本量对 Q 矩阵 CCR 的边际影响

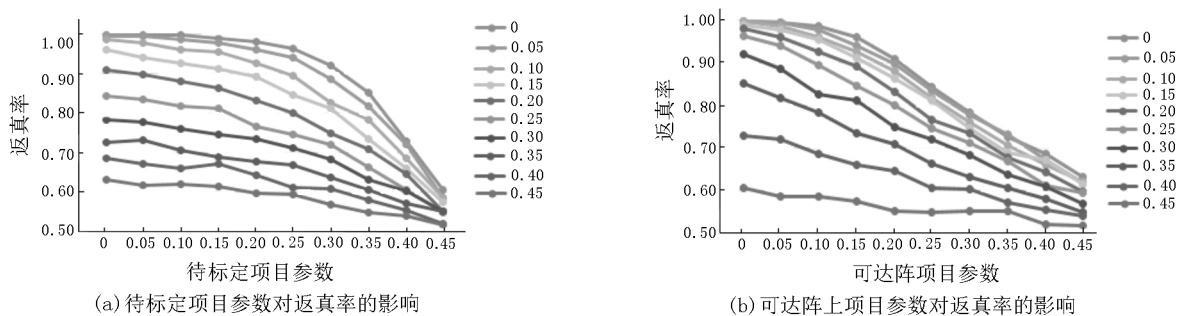


图 2 可达阵与待标定项目参数对 CCR 的影响

2.5.2 估计 Q 矩阵下知识状态的分类准确率 分类准确率可以判断该标定方法对被试者的分类是否

准确, 能否运用于实际的测验中. 即通过真实 Q 矩阵与估计 Q 矩阵之间的差异来判断该标定方法. 如

果 2 者之间的差异较小,那么便可将该标定方法运用于实际中对被试者进行分类.表 2 给出了当样本量为 150 时,真实 Q 矩阵与估计 Q 矩阵在各个可达阵参数与待标定项目参数上之差的平均值.从总体上来说,除去待标定项目的项目参数较大的情况外,

估计 Q 矩阵与真实 Q 矩阵下知识状态的分类准确率之差,随着可达阵项目的项目参数增大而逐渐增大.并且当可达阵项目的项目参数值大于 0.30 时,2 者之差均大于 0.05.

表 1 新方法标定的 Q 矩阵的 C_{CR} (样本量为 150)

可达阵 项目参数	待标定项目参数										
	0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	M
0	0.999 8	0.998 9	0.997 8	0.991 0	0.980 4	0.962 8	0.921 7	0.852 3	0.729 7	0.603 7	0.903 8
0.05	0.996 8	0.994 8	0.986 5	0.977 4	0.962 2	0.940 9	0.886 5	0.816 8	0.720 4	0.585 4	0.886 8
0.10	0.987 1	0.978 1	0.961 1	0.955 7	0.926 5	0.893 8	0.827 1	0.783 4	0.684 1	0.584 3	0.858 1
0.15	0.961 9	0.942 4	0.927 7	0.913 3	0.893 1	0.845 8	0.810 8	0.735 5	0.659 4	0.574 2	0.826 4
0.20	0.910 3	0.897 0	0.880 9	0.864 1	0.830 8	0.799 4	0.749 5	0.708 8	0.643 7	0.551 0	0.783 5
0.25	0.844 5	0.833 5	0.817 4	0.810 8	0.764 9	0.745 2	0.718 7	0.661 1	0.604 1	0.546 5	0.734 7
0.30	0.782 1	0.778 3	0.761 1	0.746 5	0.734 0	0.711 8	0.683 0	0.630 5	0.601 1	0.550 1	0.697 8
0.35	0.726 2	0.732 3	0.704 3	0.689 0	0.677 6	0.669 0	0.635 3	0.603 4	0.571 4	0.549 0	0.655 8
0.40	0.686 0	0.671 0	0.659 4	0.669 9	0.642 8	0.610 1	0.607 5	0.578 7	0.552 7	0.518 9	0.619 7
0.45	0.629 7	0.615 1	0.619 6	0.613 1	0.596 6	0.594 0	0.568 4	0.548 6	0.539 1	0.515 1	0.583 9
M	0.852 5	0.844 1	0.831 6	0.823 1	0.800 9	0.777 3	0.740 8	0.691 9	0.630 6	0.557 8	0.755 0

表 2 真实与估计 Q 矩阵分类准确率之差 (样本量为 150)

可达阵 项目参数	待标定项目参数										
	0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	M
0	0.000 0	0.000 0	0.000 0	0.000 3	0.000 6	0.000 7	0.002 0	-0.000 4	-0.020 2	-0.050 0	-0.006 7
0.05	0.000 0	0.000 1	0.000 3	0.000 9	0.004 3	0.008 0	0.009 4	0.002 4	-0.011 4	-0.042 8	-0.002 9
0.10	0.000 0	0.001 0	0.003 7	0.004 5	0.009 9	0.015 7	0.016 3	0.013 4	0.002 3	0.037 3	0.003 0
0.15	0.001 6	0.005 5	0.009 9	0.014 5	0.020 3	0.032 5	0.033 6	0.019 5	0.002 7	-0.019 2	0.012 1
0.20	0.009 2	0.017 9	0.023 2	0.033 2	0.048 8	0.056 9	0.053 7	0.029 0	0.018 1	-0.014 6	0.027 5
0.25	0.037 0	0.043 3	0.061 0	0.072 0	0.080 4	0.075 6	0.066 8	0.057 1	0.017 6	-0.003 5	0.050 7
0.30	0.092 1	0.091 6	0.109 9	0.110 2	0.113 6	0.105 4	0.093 8	0.065 6	0.036 4	0.001 4	0.082 0
0.35	0.144 1	0.149 8	0.167 2	0.165 0	0.158 6	0.136 8	0.117 9	0.081 2	0.049 8	0.009 3	0.118 0
0.40	0.204 0	0.231 7	0.219 1	0.212 8	0.190 8	0.176 1	0.138 5	0.104 0	0.057 2	0.017 5	0.155 2
0.45	0.325 4	0.325 1	0.301 0	0.270 9	0.238 2	0.200 7	0.158 1	0.122 3	0.071 6	0.025 3	0.203 9
M	0.081 3	0.086 6	0.089 5	0.088 4	0.086 6	0.080 8	0.069 0	0.049 4	0.022 4	-0.011 4	0.064 3

3 结论

该文根据 DINO 模型中可达阵 R 与简化 Q 矩阵在列向量上存在布尔“或”的关系,提出基于可达阵的补偿模型 Q 矩阵标定方法.通过模拟实验对新方法的可行性进行研究.模拟实验研究表明:

- 1) 总体上 Q 矩阵的 C_{CR} 随着样本量的增加而逐渐增大,并且与待标定项目参数影响相比,该方法所受的可达阵项目参数影响相对更大.
- 2) 当样本量在 150 且项目参数控制在 0.25 以下时,随样本量的增加 Q 矩阵的 C_{CR} 的增幅不大,样本量为 150 可视为较为理想的条件.
- 3) 当可达阵项目参数小于 0.20 且待标定的项

目参数小于 0.25 时,使用该方法得到的 Q 矩阵的 C_{CR} 达到 0.90 以上.

4) 当可达阵对应题目的失误与猜测参数都小于 0.25 时,估计 Q 矩阵与真实 Q 矩阵之间所得 A_{MR} 之差很小.

由于模拟研究使用的是 DINO 这种以 0-1 计分方式的模型,因此,以下这几个方面的问题都将是以后值得研究的方向:

- 1) 能否将这种 Q 矩阵标定方法推广到多级评分模型?
- 2) 属性层级结构 5 种类型,分别为线型、收敛型、发散型、无结构型和离散型.对于这几类属性层级结构,使用该方法能否得到较高准确率的 Q 矩阵.换言之,这 5 种属性层级结构对标定方法是否有

影响?

3) 模拟研究只研究了 5 个属性, 如果只增加属性数量, 是否会影响估计 Q 矩阵的准确率?

4) 如何减少可达阵项目参数对 Q 矩阵标定方法的影响?

4 参考文献

- [1] De la Torre J. An empirically based method of Q -matrix validation for the DINA model: development and applications [J]. Journal of Educational Measurement, 2008, 45(4): 343-362.
- [2] 涂冬波, 蔡艳, 戴海琦. 基于 DINA 模型的 Q 矩阵修正方法 [J]. 心理学报, 2012, 44(4): 558-568.
- [3] Decarlo L T. Recognizing uncertainty in the Q -matrix via a Bayesian extension of the DINA model [J]. Applied Psychological Measurement, 2012, 36(6): 447-468.
- [4] 喻晓峰, 罗照盛, 秦春影, 等. 基于作答数据的模型参数和 Q 矩阵联合估计 [J]. 心理学报, 2015, 47(2): 273-282.
- [5] Liu Jingchen, Xu Gongjun, Ying Zhiliang. Data-driven learning of Q -matrix [J]. Applied Psychological Measurement, 2012, 36(7): 548.
- [6] Chiu Chia Yi. Statistical refinement of the Q -matrix in cognitive diagnosis [J]. Applied Psychological Measurement, 2013, 37(8): 598-618.
- [7] 陈平, 辛涛. 认知诊断计算机化自适应测验中在线标定方法的开发 [J]. 心理学报, 2011, 43(6): 710-724.
- [8] 汪文义, 丁树良. 题库结构对原始题在线属性标定准确性之影响研究 [J]. 心理科学, 2012, 35(2): 452-456.
- [9] 汪文义, 丁树良, 游晓峰. 计算机化自适应诊断测验中原始题的属性标定 [J]. 心理学报, 2011, 43(8): 964-976.
- [10] Chen Ping, Xin Tao, Wang Chun, et al. Online calibration methods for the DINA model with independent attributes in CD-CAT [J]. Psychometrika, 2012, 77(2): 201-222.
- [11] Maris E. Estimating multiple classification latent class models [J]. Psychometrika, 1999, 64(2): 187-212.
- [12] Templin J L, Henson R A. Measurement of psychological disorders using cognitive diagnosis models [J]. Psychol Methods, 2006, 11(3): 287-305.
- [13] 詹沛达, 王立君, 陈飞鹏. 不同因素对认知诊断 DINO 模型诊断准确率的影响 [J]. 考试研究, 2013(4): 60-67.
- [14] 杨淑群, 蔡声镇, 丁树良, 等. 求解简化 Q 矩阵的扩张算法 [J]. 兰州大学学报: 自然科学版, 2008, 44(3): 87-91.
- [15] 丁树良, 罗芬, 汪文义, 等. Q 矩阵标定的一种简便方法 [J]. 江西师范大学学报: 自然科学版, 2018, 42(2): 130-133.
- [16] 汪文义, 宋丽红, 丁树良. 基于可达阵的一种 Q 矩阵标定方法 [J]. 心理科学, 2018, 41(4): 968-975.
- [17] Chiu Chia Yi, Douglas J. A nonparametric approach to cognitive diagnosis by proximity to ideal response patterns [J]. Journal of Classification, 2013, 30(2): 225-250.

The Method for Compensatory Model's Q -Matrix Specification Based on the Reachability Matrix

WANG Wenyi¹, WANG Teng¹, SONG Lihong², GAO Peng¹

(1. College of Computer Information Engineering, Jiangxi Normal University, Nanchang Jiangxi 330022, China;

2. Elementary Educational College, Jiangxi Normal University, Nanchang Jiangxi 330022, China)

Abstract: It is very important that the calibration method of Q -matrix in cognitive diagnosis, which directly determines the correctness of the classification of individual. The augment algorithm provides fact that any column of the reduced Q -matrix can be expressed by the columns of the reachability matrix under the logical OR operation. The purpose of this study is to propose a method for compensatory model's Q -matrix specification based on the reachability matrix. Simulation results show that the performance of the new method is promising in terms of correct classification rates of examinees' attributes.

Key words: cognitive diagnostic model; the reachability matrix; compensatory model; Q -matrix specification

(责任编辑: 冉小晓)